PROPHESY

Prophesy: A Web-based Performance Analysis and Modeling System for Parallel and Grid Applications

Valerie Taylor
Texas A&M University

Xingfu Wu, Joseph Paris
Northwestern University

Rick Stevens, Ivan Judson, Mark Hereld
Argonne National Laboratory

Xingfu Wu <uuxf@ece.northwestern.edu> http://prophesy.mcs.anl.gov

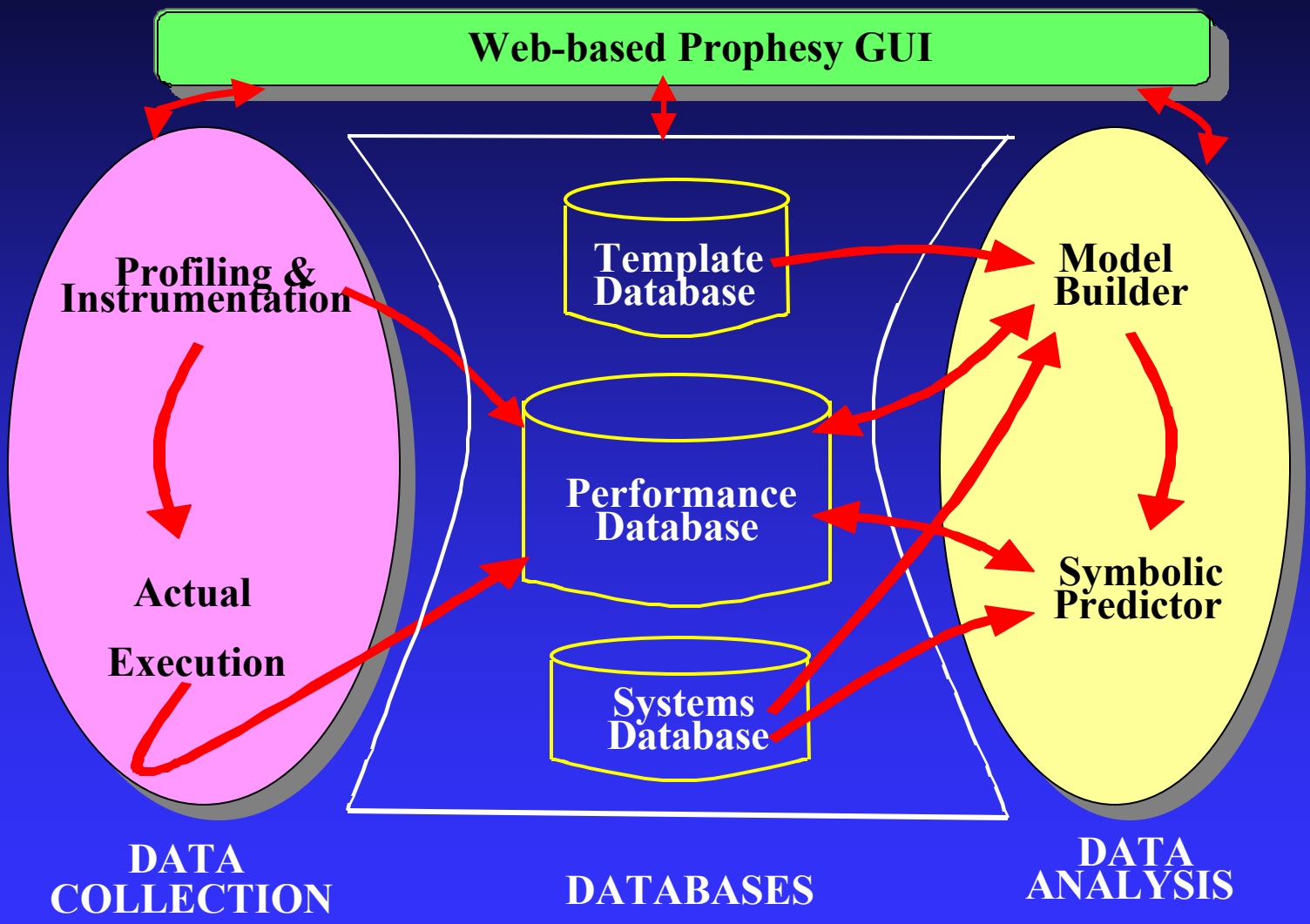
Xingfu Wu <uuxf@ece.northwestern.edu> http://prophesy.cs.tamu.edu

Outline

- Prophesy System
- Prophesy Database
- Data Collection: PAIDE System
- Data Analysis: Model Builder
- Summary

PROPHESY

Prophesy System



Xingfu Wu <uuxf@ece.northwestern.edu> http://prophesy.cs.tamu.edu

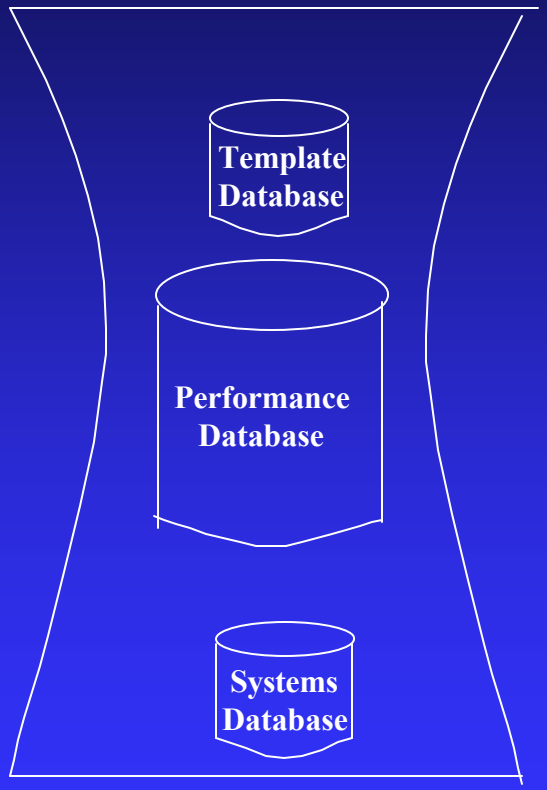
Xingfu Wu <uuxf@ece.northwestern.edu> http://prophesy.mcs.anl.gov

Prophesy System

- Data Collection
  - ✓ PAIDE system
- Prophesy Database
  - ✓ Systems Database
  - ✓ Performance Database
  - ✓ Template Database
- Data Analysis
  - ✓ Model Builder
  - ✓ Symbolic Predictor

Xingfu Wu <uuxf@ece.northwestern.edu>
http://prophesy.mcs.anl.gov

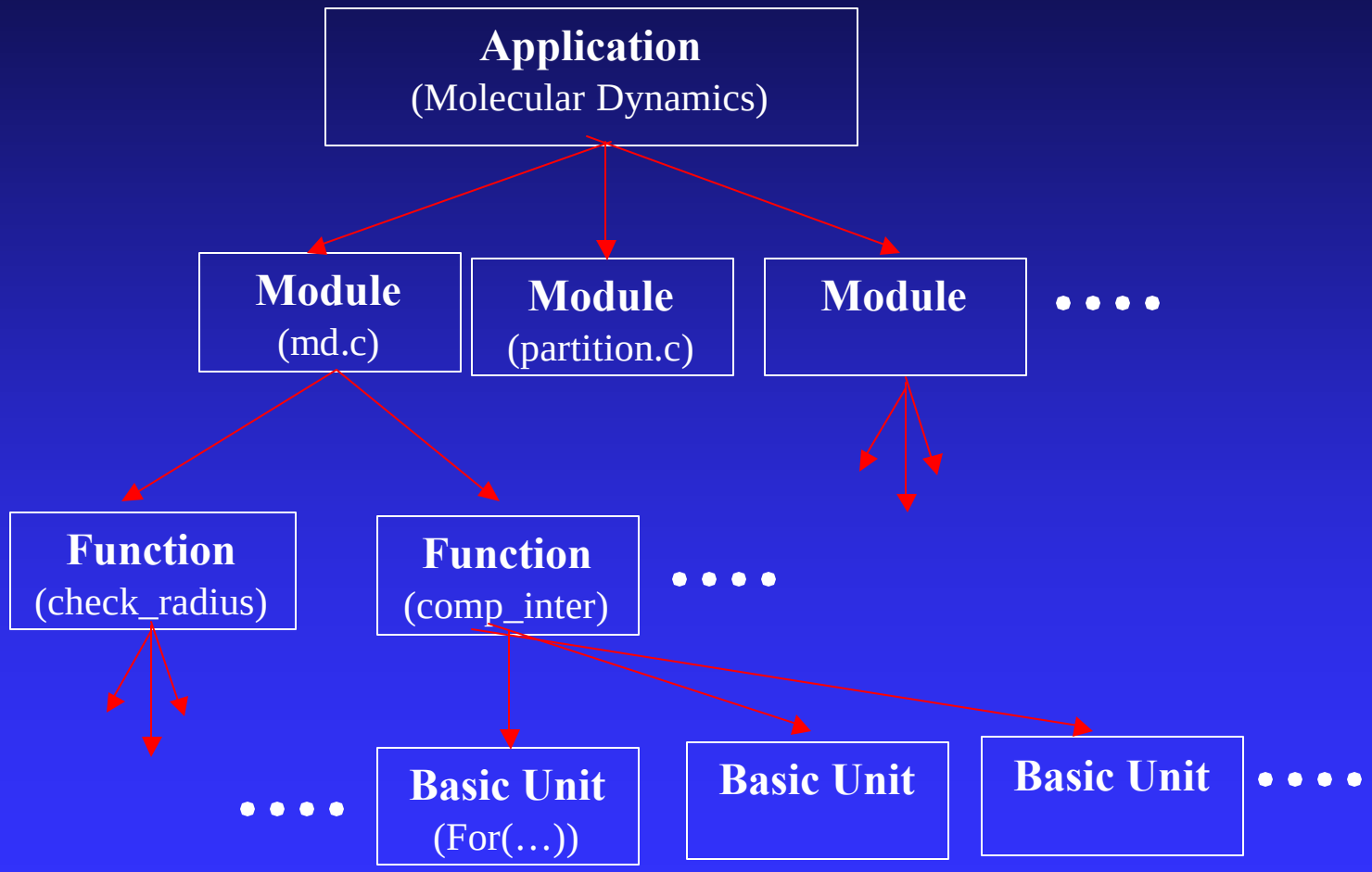
Prophesy Database



- Hierarchical organization
- Organized into 4 areas:
  - Application
  - Executable
  - Run
  - Performance Statistics

Xingfu Wu <uuxf@ece.northwestern.edu>

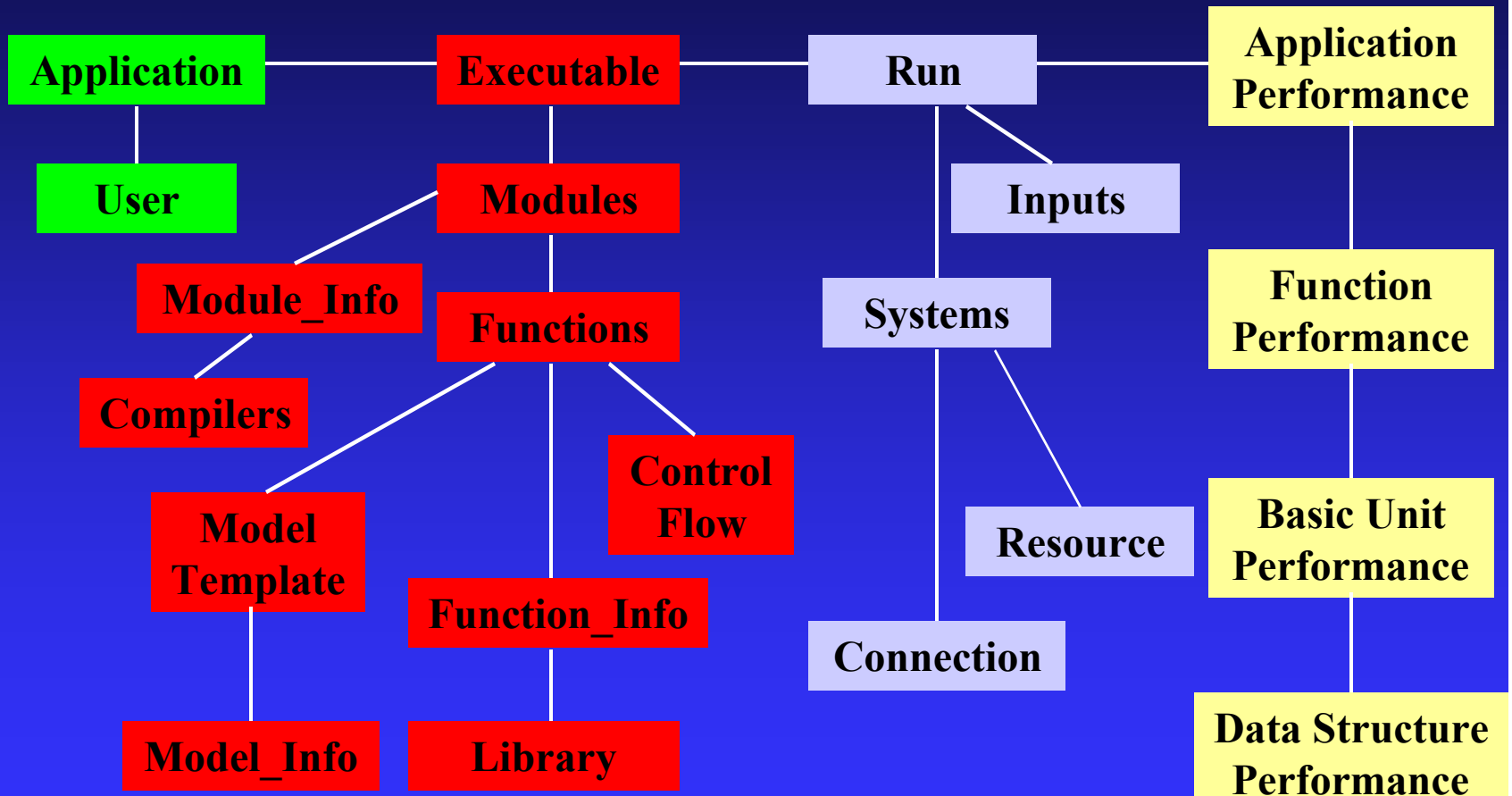
Applications



PROPHESY

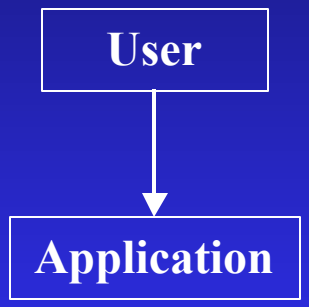
Xingfu Wu <uuxf@ece.northwestern.edu>

Prophesy Database



Xingfu Wu <uuxf@ece.northwestern.edu>
http://prophesy.cs.tamu.edu

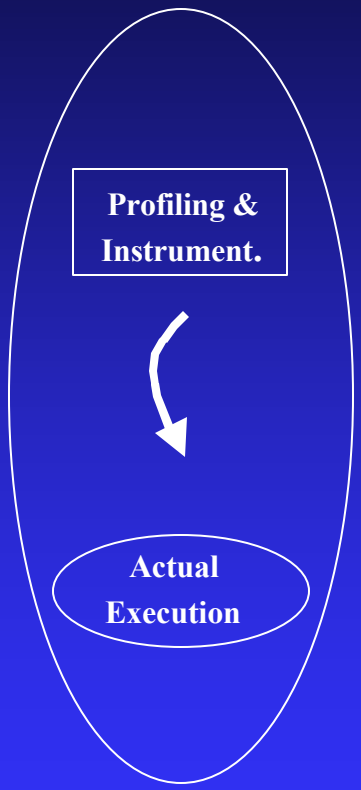
User Input



- User should register an account and an application online first.
- It requires information about the user and application such as user name, password, email, application name, and version, etc.
- Done once for all executables of the application the user owns.

Xingfu Wu <uuxf@ece.northwestern.edu> http://prophesy.cs.tamu.edu

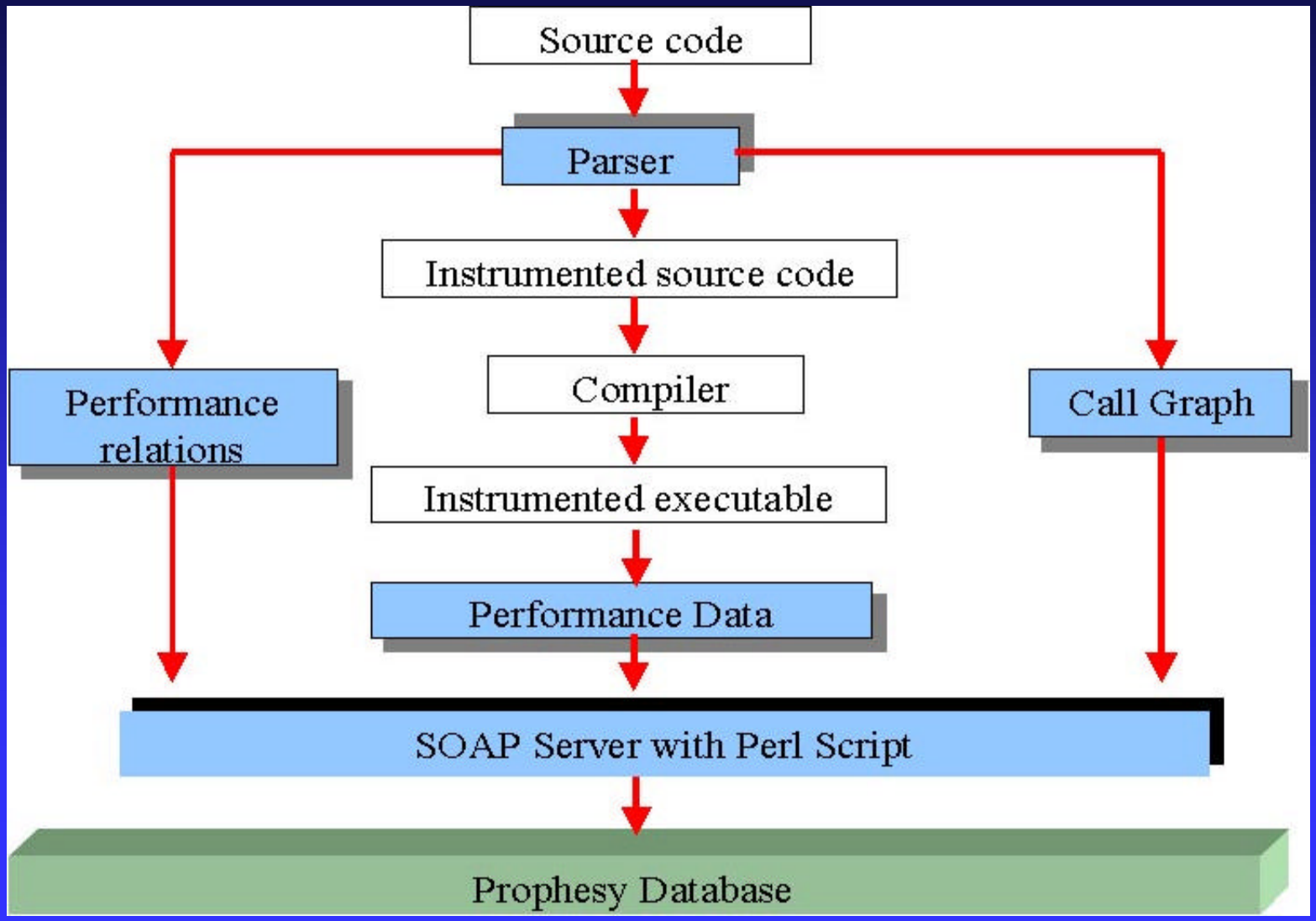
Data Collection: PAIDE System



- Automated Source-code instrumentation at the multiple levels via PAIDE
- Support for C, Fortran77 and 90 programs
- Minimal instrumentation overhead and code
- Performance Data entered into the database automatically via PAIDE or manually via web site

Xingfu Wu <uuxf@ece.northwestern.edu>

PAIDE System



Xingfu Wu <uuxf@ece.northwestern.edu>

- Options:

-ALL: Instrument all procedures and outer loops

-PROC: Instrument all procedures

-LOOP: Instrument all loops

-NOP: Instrument all procedures not nested in loops

-FTP: Use Perl SOAP scripts to automatically transfer performance data to the Prophesy database

Default: Instrument procedures and outer loops

Xingfu Wu <uuxf@ece.northwestern.edu> http://prophesy.cs.tamu.edu

- Performance Data Files

For user and application:

- ✓ User name
- ✓ Password
- ✓ Email
- ✓ Application Name
- ✓ Application Version

Xingfu Wu <uuxf@ece.northwestern.edu>

For each executable:

- ✓ Executable Name
- ✓ Problem Size
- ✓ Total Number of Processors
- ✓ Total Execution Time
- ✓ Processor Number
- ✓ System Name
- ✓ Run Date and Time

Xingfu Wu <uuxf@ece.northwestern.edu> http://prophesy.cs.tamu.edu

For each event (procedure or loop):

- Event ID
- Start Line Number
- End Line Number
- Event type (Procedure or Loop)
- Procedure Name (if event type is Procedure)
- Caller Name
- Module Name
- Runtime
- Square of runtimes

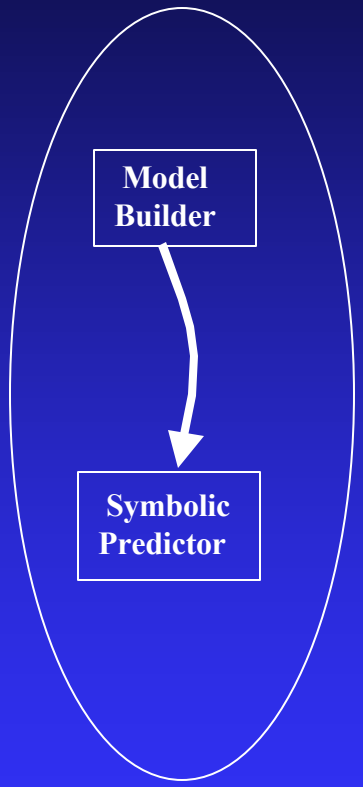
Xingfu Wu <uuxf@ece.northwestern.edu> http://prophesy.cs.tamu.edu

Data Entry

- Use Perl SOAP scripts to automatically process the performance data files at the end of program execution, and put the data into the Prophesy database.
- Use web form interfaces to manually put the data into the Prophesy database.
- Use Perl SOAP script to automatically process performance data files generated by SvPablo,and put them into the database.

Xingfu Wu <uuxf@ece.northwestern.edu>

Data Analysis: Model Builder



- Develop performance models

- Make runtime predictions
- Identify best implementation
- Identify performance trends and performance bottlenecks

Xingfu Wu <uuxf@ece.northwestern.edu>
http://prophesy.cs.tamu.edu

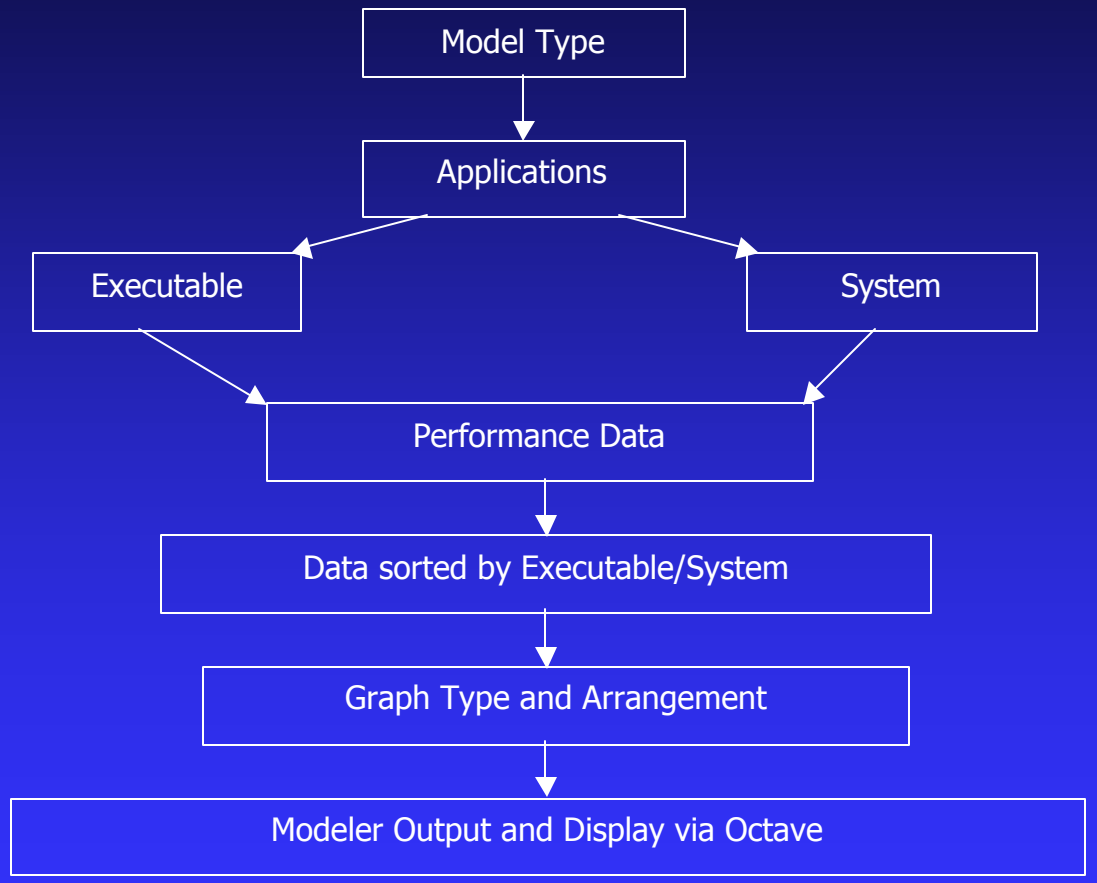
Develop Performance Models

- Utilize information in the Prophesy databases
  - ✓ Performance database
  - ✓ Template database
  - ✓ System database
- Three techniques
  - ✓ Curve Fitting
  - ✓ Parameterization
  - ✓ Coupling

PROPHESY

Xingfu Wu <uuxf@ece.northwestern.edu>

Model Builder Framework



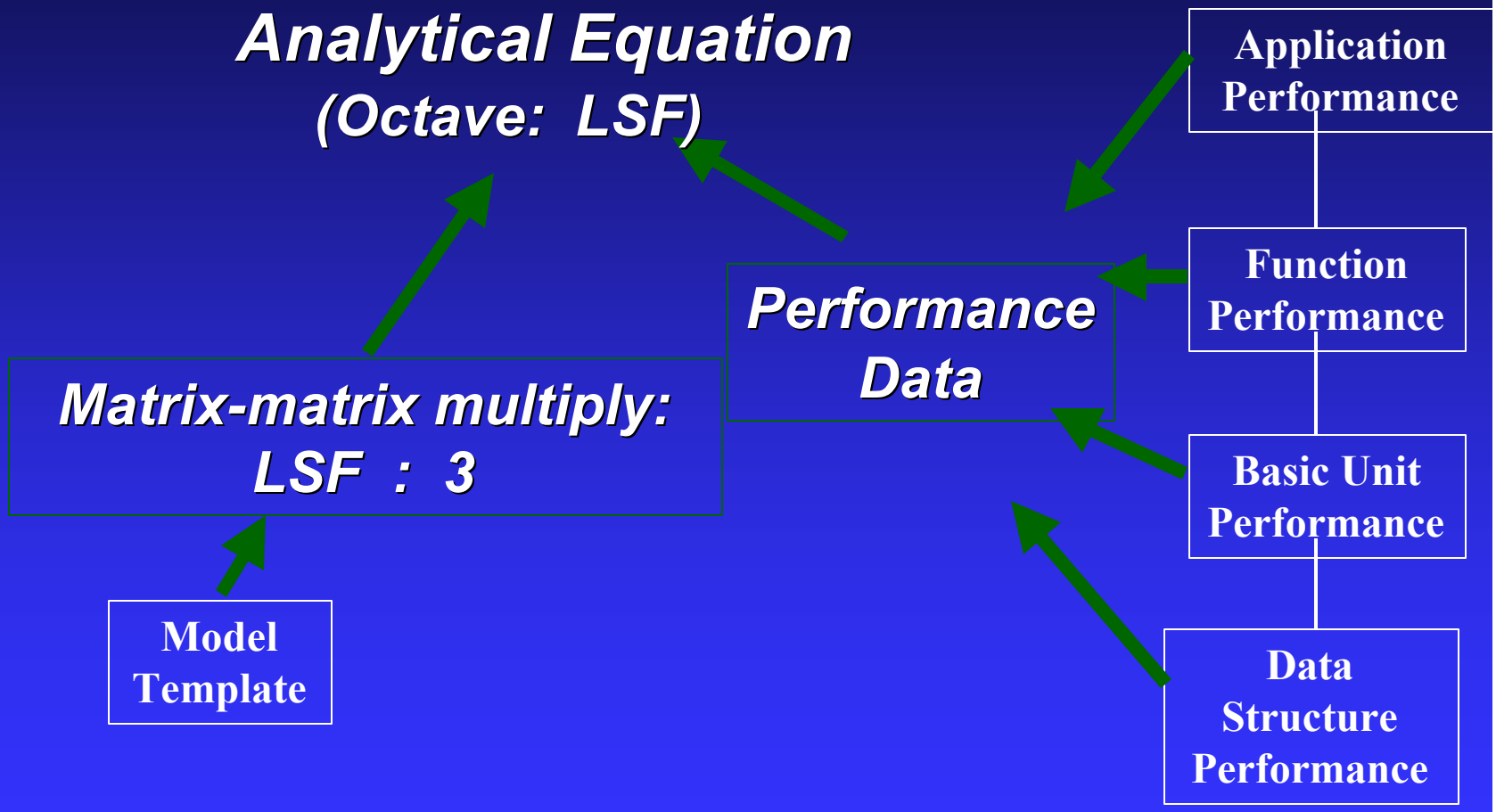
Xingfu Wu <uuxf@ece.northwestern.edu>
http://prophesy.cs.tamu.edu

Curve Fitting Method

- Uses least squares
- Uses database information
  - Executable information
    - ✓ Runtime
    - ✓ Inputs (problem size)
    - ✓ Number of Processors
    - ✓ User selected model order
- Does not expose system parameters

Xingfu Wu <uuxf@ece.northwestern.edu>

Curve Fitting: Usage



Xingfu Wu <uuxf@ece.northwestern.edu> http://prophesy.cs.tamu.edu

Parameterization Method

- Requires manual analysis of the kernel or function
  - ✓ Hand count operations
  - ✓ Expose system parameters
  - ✓ Only needs to be done once per kernel
- Uses database information
  - ✓ System database
  - ✓ Model template database

Xingfu Wu <uuxf@ece.northwestern.edu>
http://prophesy.cs.tamu.edu

Parameterization: Usage

Analytical Equation

(Octave: Parameterization)

Matrix-matrix multiply: Parameterization :

Parameter(P, SGI Origin2000, N, ADDM, MPISR, MPIBC)

Model Template

System Data:

MPISR, MPIBC, ADDM

Systems

Resource

Connection

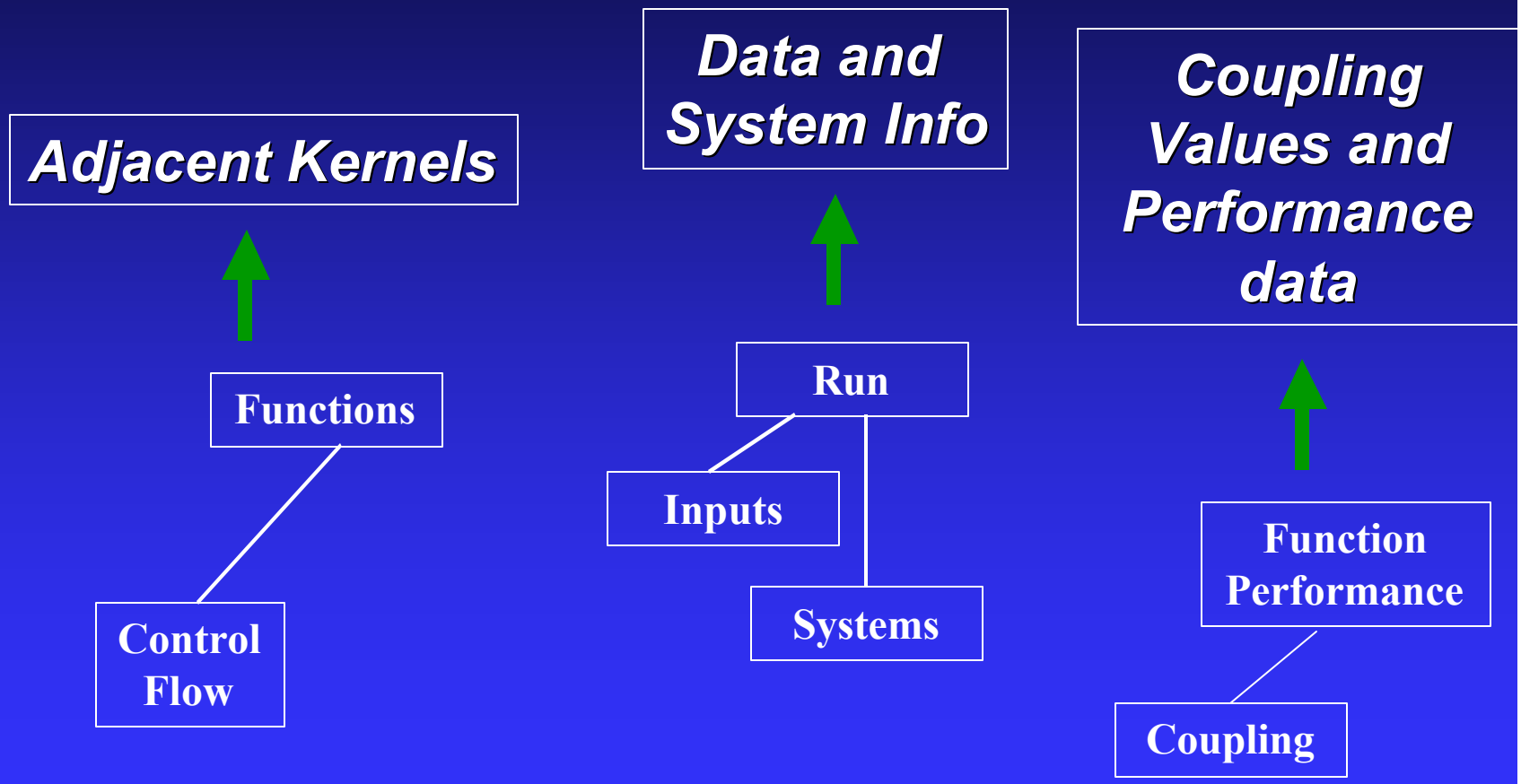
Xingfu Wu <uuxf@ece.northwestern.edu>
http://prophesy.cs.tamu.edu

Coupling Method

- Represents an application in terms of its kernels or components
- Does not require manual analysis
- Uses database information
  - ✓ Coupling values
  - ✓ Performance data

Xingfu Wu <uuxf@ece.northwestern.edu> http://prophesy.cs.tamu.edu

Coupling Method: Usage



Xingfu Wu <uuxf@ece.northwestern.edu>
http://prophesy.mcs.anl.gov

Summary

- Instrument at the level of basic blocks and/or procedures automatically via PAIDE.
- Enter data into the database automatically via PAIDE or manually via web site.
- Present the automated modeling component of Prophesy with three techniques:
  - ✓ Curve Fitting
  - ✓ Parameterization
  - ✓ Coupling

Xingfu Wu <uuxf@ece.northwestern.edu> http://prophesy.mcs.anl.gov

Prophesy System Web Page